

Running head: HOW OFTEN IS “OFTEN” REVISITED

**How Often is “Often” Revisited: The Meaning and Linearity of Vague
Quantifiers Used on the National Survey of Student Engagement**

Thomas F. Nelson Laird
Assistant Professor
Indiana University
1900 East Tenth Street
Eigenmann Hall, Suite 419
Bloomington, IN 47406-7512
tflaird@indiana.edu

Ali Korkmaz
Assistant Research Scientist
Indiana University Center for Postsecondary Research

Pu-Shih Daniel Chen
Assistant Professor
University of North Texas

Paper presented at the Annual Meeting of the American Educational Research Association,
April 15, 2008
San Diego, CA

Abstract

Drawing on responses from 26,204 first-year students and 36,263 seniors from 149 colleges and universities, this study examines the meaning of vague quantifiers (“sometimes,” “often,” and “very often”) relying on respondent identified quantities associated with their vague responses. We found that the meaning of the vague quantifiers varied from item to item, that median frequencies associated with the vague quantities were very close to linearly related, though many times a polynomial equation fits best, and that generally small differences exist between groups of students and between institutions in the meaning of the vague quantifiers.

How Often is “Often” Revisited: The Meaning and Linearity of Vague Quantifiers Used on the National Survey of Student Engagement

In survey research, we often ask respondents to judge frequency (“how often”), quantity (“how much”), and intensity (“how strongly”). The response options given for these questions are usually not exact, but are generally ordered. Examples of such response options include “sometimes,” “often”, and “very often” or “very little,” “some,” “quite a bit,” and “very much.” Many college student surveys use these types of comparative ratings to understand student activities, attitudes and judgments (see, for example, Higher Education Research Institute, 2007; National Survey of Student Engagement [NSSE], 2006a; Pace & Kuh, 1998). Yet, with the exception of Pace and Friedlander’s (1982) study nearly thirty years ago, the higher education research community has done little to understand how the use of these “vague quantifiers” may influence their findings. For instance, making a judgment about whether students are participating in an activity (e.g., asking questions in class) enough, requires that those making the judgment have an understanding of what doing the activity “often,” for example, means. Further, making comparisons between groups of students or between institutions assumes that the meaning of “very often” is consistent across the groups being compared. And, creating linear combinations of items (the process most often used to create scales from items) and running many types of analyses (e.g., correlational analyses) assumes that the response options are interval in nature (i.e., the difference between “very little” and “some” is the same as the difference between “some” and “quite a bit”). The goals of this study are to understand the meaning students participating in NSSE ascribe to the responses of “never,” “sometimes,” “often,” and “very often,” to determine if those meanings vary by survey item, student characteristics, and institutions, and to determine how close to linear these response options are.

Background

Survey response options that are not exact in quantity have been called vague quantifiers to emphasize the imprecision of their meanings (Walonick, 1994). In the 1940s, 1950s, and 1960s, several researchers attempted to specify the meanings of these vague quantifiers. For example, Hakel (1968) repeated an experiment of Simpson’s (1944) where participants were asked to rate twenty vague quantifiers and assign a meaning by its frequency. Both researchers used the median rating to rank the vague quantifiers. These researchers found considerable individual variation for many quantifiers (i.e., “often” meant different things to different respondents). Yet, in looking at median scores, clear patterns were observed. “Always” was ranked highest (with ratings of 99 and 100) while “never” was lowest (with a rating of 0 in both studies). “Sometimes” was a step above “never” in each study, though its rating varied slightly. “Very often” was a step below “always” with ratings of 88 and 87 in these studies. The rank order of the vague quantifiers was almost identical across the two studies. However, the rating of several quantifiers changed from one study to another. It was not clear whether these changes were due to 24-year gap between studies or differences between the make up of the respondents.

While work like that of Hakel (1968) and Simpson (1944) did lend credibility to the assumption that the frequently used vague quantifiers are ordinal in nature, a major weakness of that work is that it did not examine whether the meaning of the vague quantifiers could vary by context (e.g., questionnaire context, the context of an individual item, the individual respondent’s context). Rather, scholars assumed that a continuum could be created and the meaning of a vague quantifier would be a precise point on the continuum. Researchers in the late 1960s through the early 1980s rejected this notion and believed that the meanings of these words are flexible and

come from the contexts in which they are used (Bradburn & Miles, 1979; Chase, 1969; Parducci, 1968; Pepper & Prytulak, 1974; Pohl, 1981).

That the meaning of vague quantifiers is dependent on context continues to be the prevailing view among survey researchers. Studies in recent decades have explored how context can influence the meaning placed on quantifiers (Borgers et al., 2003; Schaeffer, 1991; Wanke, 2002; Wright et al., 1994).

In recent decades, vague quantifiers were put into context and researchers compared how the meaning of the same quantifier would vary from one setting to another (Borgers, Hox, & Sikkel, 2003; Mosteller & Youtz, 1990; Pepper & Prytulak, 1974; Schaeffer, 1991; Wänke, 2002; Wright, Gaskell, O’Muircheartaigh, 1994). For example, an airplane crashing “sometimes” may not mean the same thing as missing breakfast “sometimes” (Bradburn & Miles, 1979; Pepper & Prytulak, 1974). Due to these item-level differences in meaning, researchers differentiated items by their relative frequency or probability of occurrence (Bradburn & Miles, 1979; Mosteller & Youtz, 1990; Pepper & Prytulak, 1974; Wright, Gaskell, O’Muircheartaigh, 1994).

Various researchers examined whether vague quantifiers were interpreted differently by characteristics such as race, sex, education, age, occupation, and affect (Borgers et al., 2003; Bradburn & Miles, 1979; Schaeffer, 1991; Schimmack et al., 2000; Wright et al., 1994). For example, Schaeffer (1991) looked at whether using absolute and relative frequency of the feelings, boredom and excitement, were different by race, sex, education and age. She found significant differences in the meaning of the response categories based on race, education, and age, but no differences between males and females. Differences by age were also noted by Borgers et al. (2003). Wright et al. (1994) found differences by age, gender, and occupation.

Bradburn and Miles (1979) and Schimmack et al. (2000) found differences in interpretation of vague quantifiers depending on the respondent’s affect (positive or negative affect or pleasant or unpleasant emotions, respectively).

The potential for linear interpretations of vague quantifiers (i.e., whether the scale is interval in nature) was supported in studies like Hake’s (1968) that suggest that some quantifiers may be close to evenly spaced (e.g., the median frequencies assigned to “never,” “sometimes,” “about as often as not,” “often,” and “always” were 0, 29, 50, 74, and 100, respectively). However, the degree of linearity appears to depend on both the vague quantifiers used and the relative frequency of the activity. In Hake’s (1968) study, the difference in median ratings between “seldom” and “occasionally” was much smaller than the difference between “occasionally” and “often,” which may explain the problems in achieving evenly spaced responses noted by Pohl (1981), who used these vague quantifiers along with anchors of “never” and “always.” Pepper and Prytulak (1974) used the quantifiers “almost never,” “seldom,” “sometimes,” “frequently,” and “very often” and achieved near perfect linearity (based on plotting mean activity ratings by response category) for high and moderate frequency activities like shootings in a Hollywood Western or skipping breakfast, but lower frequency activities like plane crashes and earthquakes deviate from linearity.

In the field of higher education, surveys utilize vague quantifiers to learn about a wide range of behaviors, attitudes, interests, feelings, and judgments (e.g., Higher Education Research Institute, 2007; NSSE, 2006a; Pace & Kuh, 1998). While this type of quantifier is almost ubiquitous in higher education survey research, we found only one study, by Pace and Friedlander (1982), that reported on the meaning of different quantifiers. In their study, using questions from the College Student Experiences Questionnaire (CSEQ), students reported their

responses in two formats: first, while completing the standard questionnaire, using the vague quantifiers "never," "occasionally," "often," and "very often" and, at the end of the survey, using the more specific ratings, "never," "once or twice a year," "3 to 6 times a year," "1 or 2 times a month," "about once a week," and "more than once a week." Their analyses included responses from about 9000 students from 30 different institutions. As in all the studies described above, they found individual variation in the interpretation of each quantifier (i.e. one person's "occasionally" is another person's "often"). However, like other studies, there were clear modal differences between "occasionally" and "often," and between "often" and "very often," though linearity was difficult to assess given the use of a second limited response set. Their results also point to some differences in interpretation by item. For example, while only 11% of the respondents who asked others to read something they wrote "very often" indicated that they did this activity "more than once a week," the corresponding percentage was much higher (44%) for summarizing points in readings or class notes. Though individual and item-by-item variation was apparent, Pace and Friedlander observed relatively few differences in interpretation by gender, class standing (first-years and sophomores versus juniors and seniors), grades, or major and when differences were noted, there was no clear pattern (e.g., one group always scoring above the other). Further, Pace and Friedlander demonstrated only slight variation by institution using responses from 8 of the 30 institutions as examples and concluded that since the meaning of response categories did not vary from one institution to another there was no need to develop different norms for different types of institutions.

That only one such study exists for commonly used surveys of college students is cause for concern and warrant, in and of itself, for additional study. However, there are several limitations with the Pace and Friedlander (1982) study that warrant attention. First, the study was

done on a limited set of items. Additional survey items should be examined. Second, the response categories used to gauge the meaning of “occasionally,” “often,” and “very often” were quite limited (6 categories that were themselves not interval in nature), which inhibited their ability to test statistically for linearity in the response options, though such tests did not seem to be of interest to them or other researchers (e.g., Pepper & Prytulak, 1974). Third, the statistical tests for differences by student characteristics were not reported. Fourth, Pace and Friedlander did not examine differences among all of their institutions, only 8 example colleges and universities, which severely limits the claims one can make about institutional variation in the meaning of vague quantifiers based on their study.

Purpose

Relying on data from the 149 institutions that participated in the 2006 administration of NSSE, which uses vague quantifiers in a majority of its items, this study aims to determine the efficacy of using vague quantifiers on such a questionnaire. Specifically, we test: 1) how the meaning of vague quantifiers differs by survey question or item; 2) whether there is a linear relationship between the vague quantifiers and respondents' interpretations of these categories; and 3) how the meaning of vague quantifiers varies by student and institutional groups.

Data and Methods

NSSE annually collects data about the nature of the college experience from randomly selected tens of thousands of first-year and senior students at several hundred baccalaureate degree-granting colleges and universities (Kuh, 2001; 2003). The project routinely uses both paper and Web survey modes. For this study, 12 items from the core survey were repeated at the end of the Web version of the survey for students at a subgroup of institutions. Students were reminded of their original response to the item and were then asked to quantify their response by

indicating how many times they did the activity per day, week, month, academic term, or year. Students filled in the amount and selected the time frame (see Figure 1). The content of the 12 items focused on active and collaborative learning and student-faculty interaction.

<Insert Figure 1 about here>

Sample

Respondents included 26,204 first-year and 36,263 senior students who were randomly selected from 149 institutions representing a range of types (24 Doctoral, 53 Master’s, 48 Baccalaureate, and 24 “other” types of institutions; 92 private institutions). Student characteristics mirrored those of all students participating in NSSE 2006 (see NSSE, 2006b). Of the responding students, 77% were White, 64% were female, 4% part-time among first-year and 12% among seniors. For parental education, 41% of first-years and 43% of seniors were first-generation college students (no parent has baccalaureate degree). In terms of age, 6% of first-year (25 and older) and 16% of senior (28 and older) students were considered adult students.

Measures

Student responses to the 12 repeated items (see Table 1 for items) were recoded so that all responses were on a per week basis. Students could report frequency based on five options (per day, week, month, academic term, and academic year). Week was taken as the baseline and responses with other time frames were adjusted by appropriate multipliers (day = 5, month = .231, academic term = .1, .067, or .077 depending on academic calendar, and academic year = .033). Since various institutions use different calendar systems, we used three different multipliers for academic term: .067 for institutions with traditional semesters (15 weeks), .1 for those with a quarter system, and .077 for institutions on a four-one-four plan.

We created two versions of the student-faculty interaction and active and collaborative learning scales. The first version was based on the regular (vague) survey items and the second version was based on the absolute frequency items. Since these two versions do not use the same response scale, we standardized both prior to running predictive models.

Other student characteristics were collected as part of the regular NSSE survey administration. The institutional characteristics were taken from the National Center for Education Statistics’ Integrated Postsecondary Education Data System (IPEDS) and the Carnegie Classification of Institutions of Higher Education (2005 revised version).

Analyses

Analyses were run to answer each of the three specific research questions. First, we used median values to identify the meaning of each quantifier for each item. For example, the meaning of “sometimes” for a particular item would be two times per week if two times per week was the median of the distribution of quantities associated with a response of “sometimes.” We chose median values because of some skewed absolute frequency distributions. In addition, we used two methods to judge whether an activity is most appropriately thought of as being done per day, week, month, academic term, or academic year: 1) we adjusted scores so that the timeframe used resulted in median frequencies of around one or two for “sometimes” and 2) we determined the most frequent timeframe selected by respondents for each item.

Second, to test the linearity assumption of the response options, we examined the distribution of specific quantities by each vague quantifier and survey item. For each item, we regressed the median of quantity on the normal coding of the item (never = 1, sometimes=2, often=3, and very often = 4) and estimated linear and polynomial solutions using R-squared as

the indicator of which model fit best. We separated first-year and senior students in these analyses because NSSE analyses and reports always separate the two classes.

Third, we examined whether the meaning of response options varied by student and institutional characteristics. To answer this question, we first ran descriptive statistics for student and institutional characteristics for each scale. Since all of the scales were standardized, differences between groups for the two response options (vague and absolute) would be an effect-size. Then we used hierarchical linear modeling (HLM) (Raudenbush & Bryk, 2002), because our data were nested (students within institutions) and we wanted to estimate student and institutional effects on each regular (vague) and absolute scale for both student-faculty interaction and active and collaborative learning. At the student level, we controlled for class, gender, enrollment, traditional/adult, living arrangement, first-generation status, and discipline. At the institution level, we controlled for Carnegie type, institutional size, and control.

Standardizing the dependent variables meant that the unstandardized coefficients in the HLM models represented effect sizes, which, for the dichotomous independent variables, were standardized mean difference between the groups represented in the dichotomy. The larger the effect size, the more likely the difference represents performance that warrants serious discussion. As suggested by Rosenthal and Rosnow (1991), we consider an effect size of .10 or less to represent a trivial difference, between .10 and .30 small, between .30 and .50 moderate, and greater than .50 large. Because statistical significance is sensitive to sample size, effect sizes are particularly important for consideration in this study. The large number of cases used at both the student- and institution-level makes it more likely that very small differences would be statistically significant.

Results

Tables 1 and 2 show the median absolute frequency of each activity for the respondents using each vague quantifier by item for first-year and senior students. These two tables also provided the distribution of students by response option. The third row for each item gives our estimate of the most appropriate timeframe keeping the median frequency for “sometimes” as near to one or two as possible. It also shows how frequently it was selected as a base for students. Among the twelve items, for both first-year students and seniors, one item was classified as a per week activity (asking questions in class), five items (e.g. working with other students outside of class) were per month, five (e.g. making a class presentation) were per academic term, and one (participating in a community-based project) was per academic year.

These timeframes (shown in Tables 1 and 2) were based on our efforts to keep “sometimes” at a frequency near one or two. However, these timeframes were not always those most frequently picked by students (the rank compared to other timeframes is given in parentheses in Tables 1 and 2). Among first-year students, the timeframes in Table 1 was the most frequent choice for five items, while this was true for four items for senior students. For 6 items among first-year students and 5 items among senior students, the timeframes in Tables 1 and 2 were the second-choice, based on frequency. In nearly all cases, a second or third choice rating was for an adjacent category and usually had a frequency very close to that of the first choice category. For example, for asking questions in class, 44% of first-year students selected “day” as their reference while 43% selected the “week.”

As with the timeframes, the quantity of times associated with each quantifier varied by item with the exception of “never.” In both first-year and senior tables, “never” meant zero or very close to zero times for all twelve items, indicating that “never” is not a vague term for most respondents. However, the frequencies of the other quantifiers were vague and varied from item

to item. The frequency of “sometimes” for one activity (asking questions in class) was sometimes equal to the frequency of “very often” for another activity (worked with other students on projects during class). Or, in the case of first-year students, receiving prompt feedback from faculty “sometimes” is almost equal in absolute frequency to making a class presentation “often.”

Among the response options, the variability in meaning increased as the scale increased. For example, among the items that were viewed as per month, “very often” ranged from 5 times (received prompt feedback from faculty) to 13 times (tutored or taught other students) where as the meaning of “often” was close to 4 for each.

For each of the items, there is a clear distinction among vague quantifiers for both first-year and seniors. The meaning of each response option (“never,” “sometimes,” “often,” and “very often”) is very distinct from the others. There is some variation between first-year and seniors; however these variations within the items do not deviate into the meaning of other quantifiers for the other class. For example, asking questions in class “often” means 5 times per week among seniors but 6 times per week for first-year students. Among the different items, 8 out of 12 have these slight variations between first-year and seniors. In 7 of these cases, first-year students had higher frequencies associate with the vague quantifier than their senior counterparts. In most of the items, first-year students were closer to sometimes while seniors were closer to “often” but with the variation in meaning of “often” and “sometimes”, these discrepancies could diminish.

<Insert Table 1 about here>

<Insert Table 2 about here>

In the second step of our analyses, we regressed the median quantity per week on the numeric representation of the response options (1, 2, 3, and 4, respectively). We ran a linear and a second degree polynomial function to determine R-squares. All of the R-squared values were above .86 and all but three were over .90 (see Table 3), suggesting that the relationship among the response options approximates linearity. That said, based on the R-square values, a polynomial function fits slightly better than a straight line for nearly all of the items.

<Insert Table 3 about here>

With linearity demonstrated for each item with vague quantifiers, creating summative scales based on these items was less concerning (though it is important to note, given the difference in quantifier meaning from item to item, the meaning of each scale loses its connection to specific quantities when summed together as in our study). To examine differences between students and institutions, we generated scales by simply summing items together (then standardizing) for both the vaguely quantified items and the absolute frequency items. Though some argue for alternative scaling methods (e.g., Schriesheim & Castro, 1996), we wanted to examine differences between the regular (vague) and absolute scales based on the prevailing method of scale creation in the field of higher education.

Table 4 displays the average standardized scale scores for various sub-groups of respondents by student and institutional characteristics. For both vague scales, the largest difference was between first-year and senior students. For active and collaborative learning, the difference was over half a standard deviation (.56) and the difference was nearly as large for student-faculty interaction (.46). Interestingly, class differences for the absolute scales were near zero. These differences in the effect of class are noteworthy since the meanings of the vague quantifiers were not off by much by class. The difference in effect results from first-year students

giving slightly higher frequencies for “sometimes,” “often,” and “very often” compared to seniors across several items (seniors never had a median per week value more than .01 greater than first-year students, see Tables 1 and 2) coupled with the fact the first-year students were more likely to give answers of “never” or “sometimes.”

The differences in effect were much smaller for the other student characteristics, though some of these smaller differences are worthy of further exploration. Differences by major were generally smaller for the absolute scales. In active and collaborative learning scale, arts and humanities majors got more favorable results with absolute scale while education, business, and professional students got less favorable scale results with absolute scale. However, there is not that much change in student faculty interaction. While the gender and living on-campus effects were small, they reversed in direction between the vague and absolute scales. For age, the absolute active and collaborative learning scale difference was greater than the vague scale difference, but the reverse was true for student-faculty interaction. At the institutional level, the differences were small in size, in some cases it narrowed (student-faculty interaction by size) in others (active and collaborative learning by size and control) it accentuated.

<Insert Table 4 about here>

Tables 5 and 6 present the models estimating the relationships between the student and institutional characteristic and the active and collaborative learning and student-faculty interaction scales. We first ran the base (unconditional) models, then the full models were run with student and institutional characteristics added at the appropriate levels. The variance in active and collaborative learning attributable to institutions was approximately 5% for the vague scale and approximately 4% for the absolute scale. For student-faculty interaction, those figures were 5% and 3%, respectively. This small amount of between-institution variation suggests that

variability in quantifier meaning is mostly a phenomenon rooted at the individual level. That said, the slightly greater variability observed with the vague scales suggests that there is some institution-to-institution variation in the meaning of the vague quantifiers.

At the student level, the differences in effect that were apparent in Table 4 are generally apparent between the two models for both active and collaborative learning and student-faculty interaction. Large differences in the effect of class were observed for both sets of scales. Men reported significantly higher frequencies of participation for active and collaborative learning and student-faculty interaction, but the gender differences were quite small for the vague scales (and actually in the reverse direction for active and collaborative learning). Differences in the effects of major and race (particularly African Americans) were noticeable in the active and collaborative learning scales, though generally pretty small. Further investigation is warranted to determine how much this is an indication of variability in the meaning of the vague quantifiers by major and race.

<Insert Table 5 about here>

<Insert Table 6 about here>

Discussion and Conclusion

As expected, the results of this study support the notion that respondents adjust the timeframe of their responses based on the item to which they respond. The frequency associated with “sometimes” for one item can be equal to the frequency associated with “very often” for another item. Asking questions in class “sometimes” was equivalent in absolute frequency to working with other students during or outside the class “very often.”

Beyond assigning median frequencies to the vague quantifiers for each item, due to the way we asked respondents to quantify their earlier vague responses, we were able to make a

judgment about the appropriate timeframe for each activity (day, week, month, academic term, or academic year). That the items in our study have different timeframes raises questions about how best to scale such items and highlights the fact that in simply summing the items together (which may be warranted if, for example, discussing career plans with a faculty member “sometimes” is as important as asking questions in class “sometimes” for students’ development or learning), researchers lose a connection to specific amounts.

Interestingly, first-year students and seniors seem to interpret vague quantifiers a bit differently. First-year students’ “sometimes” or “often” meant more frequent activity than their senior counterparts for many items. Though for individual items these differences were relatively small, when items were combined into scales, the differences were accentuated. This suggests that the practice of NSSE staff, to run analyses and report findings separately for first-year students and seniors, is wise to avoid problems that could be caused by differences in how first-year and senior students interpret categories differently.

Our results suggest that there may be differences in interpretation by race, gender, and major as well, but these differences warrant more investigation. For example, receiving feedback from an instructor “often” might mean similar things across disciplines but working with groups during or outside of class “often” might have different interpretations for an education major or arts and humanities major. In our HLM models, the vague active and collaborative learning scale favored education majors while the absolute frequency scale favored arts and humanities majors. Yet, the differences by discipline were much smaller for student-faculty interaction. It may be the case that prevailing pedagogical approaches may influence the meaning of vague quantifiers used to denote frequency of participation in associated instructional activities (active and collaborative learning), but that activities such as student-faculty interaction may have more

common acceptance across disciplines and therefore the vague quantifications of those activities may be more commonly shared.

In most of the gender studies related to student engagement, female students are reported to be more engaged than male students (Kinzie et al., 2007). However, from this study we found that male students were more engaged when asked with absolute frequency items. Our findings are limited to two specific scales, but do raise questions about whether some of the gender differences found in the literature are artifacts of how we measure things rather than real gender differences. This is an area worth more in-depth study with a wider range of engagement scales.

Our study suggests that differences in meaning from institution to institution are generally pretty small. We found, for example, that the meaning of “often” did not change from one type of institution to another. Among the Carnegie classifications, Baccalaureate Arts and Science colleges had more favorable ratings for active and collaborative learning when the absolute scale was used. But, Baccalaureate Arts and Science colleges normally (using vague quantifiers) perform higher than other Carnegie types of institutions (NSSE, 2006b). If the more precise method were used, the differences with other Carnegie types would only be accentuated for Baccalaureate Arts and Science colleges.

Finally, the NSSE items used in this study get used often in analyses that assume these ordinal measures are, in fact, close enough to interval to proceed. Our findings show that a linear function fit the median values quite well (R-squared values over .90 for nearly all items), though, it is worth noting, a polynomial function often fit the median values slightly better. The near linearity should be quite encouraging for those who have been assuming linearity without having any empirical evidence upon which to base such an assumption!

References

- Borgers, N., Hox, J., & Sikkel, D. (2003). Response quality in survey research with children and adolescents: The effect of the labeled response options and vague quantifiers. *International Journal of Public Opinion Research*, 15(1), 83-94.
- Bradburn, N. M., & Miles, C. (1979). Vague quantifiers. *Public Opinion Quarterly*, 43(1), 92-101.
- Chase, C. I. (1969). Often is where you find it. *American Psychologist*, 24(11), 1043.
- Cliff, N. (1959). Adverbs as multipliers. *Psychological Review*, 66(1), 27-44.
- Hakel, M. D. (1968). How often is often? *American Psychologist*, 23(7), 533-534.
- Higher Education Research Institute (2007). CIRP Freshman Survey. Los Angeles, CA: University of California, Los Angeles.
- Kinzie, J., Gonyea, R., Kuh, G. D., Umbach, P., Blauch, C., & Korkmaz, A. (November, 2007). *The relationship between gender and student engagement in college*. Paper presented at the Annual Conference of the Association for the Study of Higher Education, Louisville, KY.
- Kuh, G. D. (2001). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. *Change*, 33(3), 10-17, 66.
- Kuh, G. D. (2003). What we're learning about student engagement from NSSE. *Change*, 35(2), 24-32.
- Mosteller, F. & Youtz, C. (1990). Quantifying probabilistic expressions. *Statistical Science*, 5(1), 2-34.
- National Survey of Student Engagement (2006a). The College Student Report. Bloomington, IN: Indiana University.

- National Survey of Student Engagement (2006b). Engaged learning: Fostering success for all students. Bloomington, IN: Indiana University Center for Postsecondary Research.
- Pace, C. R., & Friedlander, J. (1982). The meaning of response categories: How often is “Occasionally,” “Often,” and “Very often”? *Research in Higher Education, 17*(3), 267-281.
- Pace, C. R., & Kuh, G. D. (1998). College Student Experiences Questionnaire, fourth edition. Bloomington, IN: Indiana University.
- Pepper, S., & Prytulak, L. S. (1974). Sometimes frequently means seldom: Context effects in the interpretation of quantitative expressions. *Journal of Research in Personality, 8*, 95-101.
- Pohl, N. F. (1981). Scale considerations in using vague quantifiers. *Journal of Experimental Education, 49*(4), 235-240.
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.
- Rosenthal, R. & Rosnow, R.L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.
- Schaeffer, N. C. (1991). Hardly ever or constantly? Group comparisons using vague quantifiers. *Public Opinion Quarterly, 55*(3), 392-423.
- Schimmack, U., Oishi, S., Diener, E. & Suh, E. (2000). Facets of affective experiences: A framework for investigations of trait effect. *Personality and Social Psychology Bulletin, 26*, 655-668.
- Schriesheim, C. A., & Castro, S.L. (1996). Referent effects in the magnitude estimation scaling of frequency expressions for response anchor sets: An empirical investigation. *Educational and Psychological Measurement, 56*, 557-569.

- Simpson, R. (1944). The specific meaning of certain terms indicating differing degrees of frequency. *The Quarterly Journal of Speech*, 30, 328-330.
- Walonick, D.S. (1998). Do researchers influence survey results with their question wording choices? *Dissertation Abstract International*, 55(09), 3012. (UMI No: AAT 9504342)
Retrieved April 23, 2007, from Dissertations and Theses database.
- Wänke, M. (2002). Conversational norms and the interpretation of vague quantifiers. *Applied Cognitive Psychology*, 16, 301-307.
- Wright, D.B., Gaskell, G.D., & O’Muirheartaigh, C.A. (1994). How much is ‘Quite a bit’? Mapping between numerical values and vague quantifiers. *Applied Cognitive Psychology*, 8, 479-496.

Figure 1. Example questionnaire page soliciting absolute frequency associated with vague quantifier

The earlier question:

In your experience at your institution during the current school year, about how often have you...

Very often Often Some-times Never
▼ ▼ ▼ ▼

Asked questions in class or contributed to class discussions

Your previous response: <<response inserted here>>

Please specify the number of times you typically did this activity.
Enter a number (e.g., 1,2,3) and indicate a unit of time (e.g, day, week, month):

Times(s) Per

- Day
- Week
- Month
- Academic term
- Academic year

Table 1. First-year students’ responses converted to times and items’ contextual translation

Items	Context	Never	Sometimes	Often	Very Often
Asked questions in class	% selected	3	35	36	27
	Median per Week	0.07	2.00	6.00	15.00
	<i>Week (2nd)</i>	0.07	2	6	15
Made a class presentation	% selected	13	54	26	7
	Median per Week	0.00	0.20	0.34	0.46
	<i>Academic Term (1st)</i>	0	2	3	5
Worked with other students on projects during class	% selected	12	47	31	9
	Median per Week	0.00	0.34	1.00	2.00
	<i>Month (2nd)</i>	0	1	4	9
Worked with classmates outside of class	% selected	11	44	32	13
	Median per Week	0.00	0.23	1.00	2.00
	<i>Month (2nd)</i>	0	1	4	9
Tutored or taught other students	% selected	49	36	11	5
	Median per Week	0.00	0.23	1.16	3.00
	<i>Month (3rd)</i>	0	1	5	13
Participated in a community-based project	% selected	64	24	8	3
	Median per Week	0.00	0.10	0.46	1.00
	<i>Academic Year (1st)</i>	0	1	7	15
Discussed ideas from your readings or classes with others outside of class	% selected	5	35	37	23
	Median per Week	0.00	0.69	2.00	5.00
	<i>Month (2nd)</i>	0	3	9	22
Discussed grades or assignments with an instructor	% selected	7	43	32	18
	Median per Week	0.00	0.20	0.46	0.92
	<i>Academic Term (1st)</i>	0	2	5	9
Talked about career plans with a faculty member	% selected	24	47	20	9
	Median per Week	0.00	0.13	0.23	0.46
	<i>Academic Term (1st)</i>	0	1	2	5
Discussed ideas from your readings with faculty members outside of class	% selected	40	41	14	6
	Median per Week	0.00	0.23	1.00	1.62
	<i>Academic Term (2nd)</i>	0	2	10	16
Received prompt feedback from faculty	% selected	6	37	41	16
	Median per Week	0.00	0.33	1.00	1.39
	<i>Month (1st)</i>	0	1	4	6
Worked with faculty members on activities other than coursework	% selected	58	28	10	4
	Median per Week	0.00	0.13	0.69	1.00
	<i>Academic Term (2nd)</i>	0	1	7	10

Table 2. Senior students’ responses converted to times and items’ contextual translation

Items	Context	Never	Sometimes	Often	Very Often
Asked questions in class	% selected	1	24	33	42
	Median per Week	0.07	2.00	5.00	15.00
	<i>Week (2nd)</i>	0.07	2	5	15
Made a class presentation	% selected	4	33	39	25
	Median per Week	0.00	0.13	0.27	0.46
	<i>Academic Term (1st)</i>	0	1	3	5
Worked with other students on projects during class	% selected	10	44	31	15
	Median per Week	0.00	0.23	1.00	2.00
	<i>Month (3rd)</i>	0	1	4	9
Worked with classmates outside of class	% selected	5	35	35	25
	Median per Week	0.00	0.23	1.00	2.00
	<i>Month (3rd)</i>	0	1	4	9
Tutored or taught other students	% selected	38	37	14	11
	Median per Week	0.00	0.23	1.00	3.00
	<i>Month (3rd)</i>	0	1	4	13
Participated in a community-based project	% selected	51	31	12	6
	Median per Week	0.00	0.07	0.40	1.00
	<i>Academic Year (1st)</i>	0	1	6	15
Discussed ideas from your readings or classes with others outside of class	% selected	3	31	38	28
	Median per Week	0.00	0.69	2.00	5.00
	<i>Month (2nd)</i>	0	3	9	22
Discussed grades or assignments with an instructor	% selected	4	35	34	27
	Median per Week	0.00	0.20	0.46	0.69
	<i>Academic Term (1st)</i>	0	2	5	7
Talked about career plans with a faculty member	% selected	13	40	27	19
	Median per Week	0.00	0.10	0.27	0.47
	<i>Academic Term (1st)</i>	0	1	3	5
Discussed ideas from your readings with faculty members outside of class	% selected	25	46	18	10
	Median per Week	0.00	0.23	1.00	1.62
	<i>Academic Term (2nd)</i>	0	2	10	16
Received prompt feedback from faculty	% selected	3	29	47	21
	Median per Week	0.00	0.23	1.00	1.16
	<i>Month (2nd)</i>	0	1	4	5
Worked with faculty members on activities other than coursework	% selected	41	34	16	9
	Median per Week	0.00	0.13	0.54	1.00
	<i>Academic Term (2nd)</i>	0	1	5	10

Table 3. R-square for linearity and polynomial assumptions for senior students’ responses

Items	Assumption	First-year R-square	Senior R-square
<i>Active and Collaborative Learning Items</i>			
Asked questions in class or contributed to class discussions	Linear	.90	.86
	Polynomial	.99	.98
Made a class presentation	Linear	.99	.99
	Polynomial	.99	.99
Worked with other students on projects during class	Linear	.95	.94
	Polynomial	.99	.99
Worked with classmates outside of class to prepare class assignments	Linear	.94	.94
	Polynomial	.99	.99
Tutored or taught other students (paid or voluntary)	Linear	.88	.86
	Polynomial	.99	.99
Participated in a community-based project (e.g., service learning) as part of a regular course	Linear	.92	.89
	Polynomial	.99	.99
Discussed ideas from your readings or classes with others outside of class (students, family members, co-workers, etc.)	Linear	.91	.91
	Polynomial	.99	.99
<i>Student-Faculty Interaction Items</i>			
Discussed grades or assignments with an instructor	Linear	.96	.99
	Polynomial	.99	.99
Talked about career plans with a faculty member or advisor	Linear	.97	.98
	Polynomial	.98	.99
Discussed ideas from your readings or classes with faculty members outside of class	Linear	.96	.96
	Polynomial	.98	.99
Received prompt written or oral feedback from faculty on your academic performance	Linear	.98	.93
	Polynomial	.98	.93
Worked with faculty members on activities other than coursework (committees, orientation, student life activities, etc.)	Linear	.95	.95
	Polynomial	.97	.99

Table 4. Standardized scores with regular (vague) and absolute measurements for student and institutional characteristics

Categories	Values	Active and Collaborative (regular)	Active and Collaborative (absolute)	Student-Faculty Interaction (regular)	Student-Faculty Interaction (absolute)
Class	First-year	-.33	-.01	-.27	-.02
	Senior	.23	.01	.19	.01
Gender	Female	.02	-.03	.00	-.04
	Male	-.04	.06	-.01	.08
Enrollment	Part-time	-.16	-.24	-.22	-.27
	Full-time	.02	.02	.02	.03
Age	Traditional	-.01	.00	.01	.02
	Adult	.06	-.01	-.06	-.14
First Generation College Student	Not First Gen	.02	.04	.02	.02
	First Gen	-.02	-.05	-.04	-.03
Live	Off-campus	.08	-.03	.03	-.03
	On-campus	-.12	.04	-.04	.04
Major	Arts & Humanities	-.06	.14	.08	.09
	Biological & Physical Sci	-.06	-.01	.07	.07
	Business & Professional	-.01	-.14	-.11	-.11
	Education	.24	.11	.04	.06
	Engineering	-.01	-.02	-.15	.00
	Social Science	-.03	.02	.06	-.01
	Other Major	-.08	-.07	-.06	-.02
Race	African American	.03	-.07	.03	-.03
	Asian American	-.07	-.09	-.08	-.01
	White	.00	.01	.00	.00
	Hispanic	.02	-.02	-.04	-.03
	Other Ethnicity	-.13	-.16	-.16	-.13
Carnegie	Doctoral High Research	-.06	-.04	-.13	-.04
	Doctoral	-.01	-.07	-.03	-.04
	Masters Large programs	.06	.05	.09	.05
	Masters Med & Small	.02	.13	.12	.07
	Baccalaureate- Arts Sci	.01	.00	.06	.09
	Baccalaureate - Diverse	.23	.26	.09	.14
	Other Carnegie	.08	.13	.13	.12
Size	Small (<2500)	.07	.03	.09	.02
	Medium (2501-5000)	-.07	-.07	-.05	-.04
	Large (5001-10000)	-.07	-.09	-.15	-.10
	Very Large (> 10000)	-.09	-.12	-.11	-.07
Control	Public	.08	.10	.10	.06
	Private	.03	-.07	.03	-.03

Table 5. HLM coefficients for active and collaborative scales

		Active & Collaborative (Regular)			Active & Collaborative (Absolute)		
		Base	Final		Base	Final	
			Coef (SE)	Sig.		Coef (SE)	Sig.
	Intercept	-.05(.03)	-.50 (.12)	***	.01(.02)	-.15(.10)	
Carnegie Doctoral as reference	Doc High Research		-.09 (.07)			-.13(.06)	*
	Masters Large		-.09 (.05)			-.11(.05)	*
	Masters Med & Sm		.04 (.07)			.00(.05)	
	Bac-Arts & Sci		-.01 (.07)			.07(.06)	
	Bac-Diverse		-.09 (.14)			-.03(.08)	
	Other Carnegie		.12 (.06)	*		.14(.07)	*
	Size		-.05 (.03)			-.04(.02)	*
	Private		.13 (.06)	*		.12(.05)	*
	Senior		.60 (.02)	***		.06(.02)	*
	Male		-.04 (.01)	**		.09(.02)	***
	Fulltime		.26 (.03)	***		.23(.03)	***
	Adult		.03 (.02)			.12(.03)	***
	First-generation		-.03 (.01)	*		-.03(.01)	*
	On-campus		.02 (.02)			.01(.02)	
Major Biological & Physical Science as reference	Arts & Humanities		.01 (.02)			.14(.03)	***
	Business & Prof		.10 (.03)	**		-.08(.03)	**
	Education		.27 (.03)	***		.15(.03)	***
	Engineering		.09 (.05)	*		.00(.04)	
	Social Science		-.03 (.03)			.02(.02)	
	Other Major		-.02 (.03)			-.07(.02)	**
Race White as reference	African American		.13 (.04)	**		.02(.03)	
	Asian American		-.07 (.02)	**		-.08(.02)	**
	Hispanic		.04 (.03)			-.01(.03)	
	Other Ethnicity		.08 (.05)			.03(.05)	
<i>Variance Components</i>							
Total variance		1.04	.91		1.00	.96	
Variance within institutions		.98	.89		.96	.95	
Variance between institutions		.05	.02		.04	.01	
Proportion between institutions		4.87%			3.83%		
Variance between explained			57.06%			66.54%	
Variance within explained			9.74%			1.13%	

Table 6. HLM coefficients for active and collaborative scales

		Student-Faculty Interaction (Regular)			Student-Faculty Interaction (Absolute)		
		Base	Final	Sign.	Base	Final	Sign.
		Coef (SE)			Coef (SE)		
	Intercept	-.01(.03)	-.22(.09)	*	.04(.02)	.00(.09)	
Carnegie Doctoral as reference	Doc High Research		-.08(.05)			-.08(.05)	
	Masters Large		-.07(.05)			-.05(.04)	
	Masters Med & Sm		.03(.07)			.00(.04)	
	Bac-Arts & Sci		.07(.07)			.03(.06)	
	Bac-Diverse		-.01(.11)			.07(.06)	
	Other Carnegie		.21(.10)	*		.13(.06)	*
	Size		-.08(.02)	**		-.06(.02)	**
	Private		.06(.04)			.02(.05)	
	Senior		.48(.02)	***		.10(.02)	***
	Male		.00(.01)			.12(.02)	***
	Fulltime		.22(.02)	***		.20(.02)	***
	Adult		-.03(.02)			-.04(.03)	
	First-generation		-.01(.01)			-.03(.02)	
	On-campus		.02(.02)			.03(.01)	
Major Biological & Physical Science as reference	Arts & Humanities		.08(.03)	**		.07(.03)	**
	Business & Prof		-.07(.03)	*		-.12(.03)	***
	Education		.03(.02)			.02(.03)	
	Engineering		-.13(.04)	**		-.09(.03)	**
	Social Science		.00(.02)			-.05(.02)	*
	Other Major		-.09(.03)	**		-.07(.03)	*
Race White as reference	African American		.12(.03)	***		.07(.04)	
	Asian American		-.01(.03)			-.03(.03)	
	Hispanic		.03(.04)			.04(.03)	
	Other Ethnicity		.11(.06)			.08(.05)	
<i>Variance Components</i>							
Total variance		1.00	.91		1.04	1.01	
Variance within institutions		.95	.89		1.01	1.00	
Variance between institutions		.05	.02		.03	.01	
Proportion between institutions		5.00%			2.87%		
Variance between explained			64.03%			56.46%	
Variance within explained			5.71%			1.29%	